Soil moisture retrieval from ASAR measurements over natural surfaces with a large roughness variability

G. Satalino, F. Mattia, G. Pasquariello, L. Dente
CNR – ISSIA
Via Amendola 122/D
I-70126, Bari, Italy
email: satalino@ba.issia.cnr.it

Abstract—**In this work, the accuracy of soil moisture retrieved from ASAR data over bare or sparsely vegetated surfaces is investigated by means of a simulation study. The soil moisture retrieval method is based on an optimization algorithm that appropriately inverts theoretical direct models by assimilating a priori information on surface parameters. In order to account for a large variability of roughness conditions, two complementary models have been used, namely the Integral Equation Method model and the Geometrical Optics model. The performance of the inversion method has been assessed on simulated noisy ASAR data, as a function of different a priori information quality level.**

Soil moisture; SAR; IEM model; GO model; Inversion method;

I. INTRODUCTION

The dependence of SAR backscattering on soil moisture and surface roughness, over bare fields, has been widely studied and modeled [1]. Many experiments confirmed a good sensitivity of radar measurements to soil moisture conditions. These observations make sensible to address the inverse problem, i.e. retrieving soil moisture from backscattering data. However, the experience gained with single-parameter SAR systems, such as ESR-1/2, has shown that soil moisture estimates are affected by large errors [2]. This is mainly because many combinations of soil moisture and soil roughness parameters correspond to the same measured backscattering value. In this respect, improvements of soil moisture estimation accuracy can be obtained principally by using multi-parameter radar measurements and a priori information on surface parameters. The accuracy of retrieved soil moisture also depends on the inversion algorithm adopted. Algorithms based on the inversion of theoretical models are preferred to empirical relationships because the latter are valid for limited and, often, site-dependent data sets. Unlike, model based algorithm can be adapted to a wider range of surface conditions. In this study, the Integral Equation Method (IEM) and the Geometrical Optics (GO) models are employed, because they are applicable to roughness conditions going from smooth to very rough ones [3].

Another factor that can be exploited to improve the soil moisture accuracy is the quality level of a priori information. The a priori information consists of tentative knowledge about the soil roughness and soil moisture conditions. Prior information on soil moisture content may come, for example, from networks of ground stations or form land process models. Whereas, initial guess values on surface roughness may be gathered from site agricultural calendar or from simple empirical relationships between SAR backscattering and roughness state. The a priori information is of high quality level when it is close enough to the true solution. When this occurs, more accurate estimates of soil moisture are possible by constraining the optimization algorithm. This corresponds to search solutions close to the a priori information, instead of searching within the whole parameter space.

In this study, following a methodology as in [4], a Bayesian framework is adopted to derive an optimization algorithm, which estimates soil moisture content from HH&VV SAR backscattering values at C-band. The algorithm inverts IEM and GO models appropriately assimilating a priori information on surface parameters.

In section II, the soil moisture retrieval problem and the optimization algorithm are shortly introduced. Then, the simulated data set is presented. Subsequently, the dependence of the retrieved soil moisture accuracy on the simulated SAR data configurations, measurement errors and different quality level of a priori information is investigated. Finally, conclusions are given.

II. THE SOIL MOISTURE RETRIEVAL BY MODEL INVERSION

The relationship between multi-dimensional (i.e. multi-polarization, multi-angle, etc.) radar backscattering coefficient, $\boldsymbol{\sigma_0}^t$, and surface geophysical parameters, $\boldsymbol{p}$, can be modeled as:

$$\boldsymbol{\sigma_0}^t = \boldsymbol{f}(\boldsymbol{p}), \qquad (1)$$

where $\boldsymbol{f}(\cdot)$ is a direct theoretical model. In this paper, bare or sparsely vegetated surfaces only will be considered. Then, direct theoretical models are surface scattering models, namely IEM and GO.

The geophysical parameters which characterize the soil surface are: the complex soil dielectric constant ε^r, the

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
25 JUL 2005	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Soil moisture retrieval from ASAR measurementsover natural surfaces with a large roughness variability	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **CNR ISSIA Via Amendola 122/D I-70126, Bari, Italy**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001850, 2005 IEEE International Geoscience and Remote Sensing Symposium Proceedings (25th) (IGARSS 2005) Held in Seoul, Korea on 25-29 July 2005.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**4**	

roughness parameters expressed in terms of the root mean square (rms) height of soil surface s, and the correlation length l, associated to the soil surface autocorrelation function (ACF). The ACF is assumed exponential when IEM surface scattering model is used. Whereas it is Gaussian when the GO model is employed. In this case, the roughness parameter, which characterizes the very rough surfaces, is the surface rms slope m (i.e. $m = \sqrt{2}\, s / l$) instead of independently s and l.

For agricultural soils, the imaginary part of the soil dielectric constant $I(\varepsilon^r)$, is usually much smaller than the real part $R(\varepsilon^r)$. Therefore, to reduce the number of independent soil parameters, we have approximated the imaginary part of the soil dielectric constant as $I(\varepsilon^r)=R(\varepsilon^r)/10$. In the inverse problem, the volumetric soil moisture content m_v, corresponding to a soil layer about 5 cm thick, is obtained first by estimating the soil dielectric constant and then by inverting the empirical expression derived by Hallikainen et al. [5].

In this context, our goal is to find the "best" estimate of soil moisture and soil roughness $\boldsymbol{p}$ which inverts (1), given HH&VV SAR measurements $\boldsymbol{\sigma_0}^m$ and prior estimates of soil parameters $\boldsymbol{p}^0$. The SAR observations $\boldsymbol{\sigma_0}^m$ differ from the theoretical backscatter $\boldsymbol{\sigma_0}^t$ due to measurement errors. In addition, the non-linear relationship (1) between predicted SAR backscatter $\boldsymbol{\sigma_0}^t$ and geophysical parameters $\boldsymbol{p}$ may also be affected by model errors. Using the Bayes's theorem, the probability density function (pdf) of $\boldsymbol{p}$ conditioned by the observed $\boldsymbol{\sigma_0}^m$ is

$$P(\boldsymbol{p}|\, \boldsymbol{\sigma_0}^m) \propto P(\boldsymbol{\sigma_0}^m|\boldsymbol{p})\, P(\boldsymbol{p}|\, \boldsymbol{p}^0), \qquad (2)$$

where $P(\boldsymbol{\sigma_0}^m|\boldsymbol{p})$ is the pdf of $\boldsymbol{\sigma_0}^m$, knowing that the geophysical parameters are $\boldsymbol{p}$, and $P(\boldsymbol{p}|\boldsymbol{p}^0)$ is the pdf of $\boldsymbol{p}$, given the knowledge of its prior estimate $\boldsymbol{p}^0$. To obtain the "best" estimate of $\boldsymbol{p}$ either the mean or the mode of $P(\boldsymbol{p}|\boldsymbol{\sigma_0}^m)$, which are the minimum variance and the maximum likelihood estimates of $\boldsymbol{p}$, can be used, respectively. It is worth mentioning that if the pdf in (2) are Gaussian, the minimum variance and the maximum likelihood estimators coincide.

Under the hypothesis that: SAR observations $\boldsymbol{\sigma_0}^m$ and parameters $\boldsymbol{p}$ are related by

$$\boldsymbol{\sigma_0}^m = \boldsymbol{f}(\boldsymbol{p}) + \boldsymbol{\varepsilon}^{m\phi}, \qquad (3)$$

where $\boldsymbol{\varepsilon}^{m\phi}$ is a zero mean Gaussian uncorrelated vector, representing the cumulative effect of measurement and model errors; the geophysical parameters $\boldsymbol{p}$ distribute around a guess value $\boldsymbol{p}^0$, i.e.

$$\boldsymbol{p} = \boldsymbol{p}^0 + \boldsymbol{\varepsilon}^p, \qquad (4)$$

where $\boldsymbol{\varepsilon}^p$ is a zero mean Gaussian uncorrelated error. Then, the Bayesian maximum likelihood estimator of $\boldsymbol{p}$ for a generic observed $\boldsymbol{\sigma_0}^m$ and subject to prior estimate $\boldsymbol{p}^0$ is obtained by maximizing $P(\boldsymbol{p}|\boldsymbol{\sigma_0}^m)$ or equivalently by minimizing the functional

$$J(\boldsymbol{p})=1/2(\boldsymbol{\sigma_0}^m-\boldsymbol{f}(\boldsymbol{p}))^T\, \mathbf{S}^{-1}\, (\boldsymbol{\sigma_0}^m-\boldsymbol{f}(\boldsymbol{p})) + 1/2(\boldsymbol{p}-\boldsymbol{p}^0)^T\, \mathbf{G}^{-1}\, (\boldsymbol{p}-\boldsymbol{p}^0), \qquad (5)$$

where $\mathbf{S}$ and $\mathbf{G}$ are diagonal covariance matrices of $\boldsymbol{\sigma_0}^m$ and $\boldsymbol{p}$, respectively [4]. An analytical solution which minimize $J(\boldsymbol{p})$ does not exist, then an iterative approach has to be sought. An efficient and quite accurate algorithm exploited to minimize $J(\boldsymbol{p})$ is the Generalized Reduced Gradient Method [6]. When no prior information on the geophysical parameters is available, the pdf $P(\boldsymbol{p}|\boldsymbol{p}^0)$ in (2) may be modeled as a constant term, which can be dropped (i.e. it is incorporated into the normalization constant). Then the functional $J(\boldsymbol{p})$ in (5) reduces to its first term only, which represents the traditional minimization of the mean square errors between measured and predicted backscattering values, i.e.

$$J(\boldsymbol{p})=1/2(\boldsymbol{\sigma_0}^m-\boldsymbol{f}(\boldsymbol{p}))^T\, \mathbf{S}^{-1}\, (\boldsymbol{\sigma_0}^m-\boldsymbol{f}(\boldsymbol{p})). \qquad (6)$$

III. DEPENDENCE OF THE RETRIEVED SOIL MOISTURE ACCURACY ON MEASUREMENTS AND A PRIORI INFORMATION ERRORS

In this section, a simulation study on the impact of measurement, model and a priori information errors on the accuracy of soil moisture and soil roughness parameters retrieved using C-band SAR data is carried out. The attention has been focused on the ENVISAT ASAR configuration in which both VV & HH polarizations are acquired at 23° incidence angle. For the sake of comparison, some plots include both the traditional ERS SAR configuration (i.e. VV at 23° incidence angle) and a virtual configuration corresponding to HH & VV polarizations acquired contemporary at 23° and 45° incidence angles. SAR measurements have been simulated by the IEM and GO models. Since, GO predictions at HH and VV polarizations are equal, VV polarization is considered only. The selected model input parameters describe a wide range of roughness and moisture conditions, i.e. from smooth to very rough surfaces and from extremely dry to wet soils.

The IEM data were simulated by using the following intervals of values: relative dielectric constant ε^r = [3.0+j0.3, 20.0+j2.0], corresponding approximately to a volumetric moisture content interval of $m_v\%$ = [3% cm^3 cm^{-3}, 38% cm^3 cm^{-3}]; roughness vertical standard deviation s = [0.6 cm, 2.1 cm]; Exponential Auto Correlation Function (ACF) with correlation length l = [6 cm, 24 cm]. The GO data were simulated by using the same range for ε^r, but higher values for the roughness s. The range of the slope m used is the interval [0.2, 0.3], $s > 3.0$.

To quantitatively study how measurement, model and a priori information errors propagate on the retrieved parameter accuracy, noisy data have been simulated. For the sake of simplicity measurement and model errors have been represented as a unique total error, i.e. $\boldsymbol{\varepsilon}^{m\phi}$ in (3). Then, the noisy data have been obtained by adding on the theoretical IEM and GO backscattering values, a zero mean Gaussian random noise with increasing standard deviation, denoted as σ^{ε}, which ranges between 0.5 dB and 2.0 dB.

The error on a priori information, i.e. $\boldsymbol{\varepsilon}^p$ in (4), has been simulated as a Gaussian uncorrelated noise, and it has been added to the true parameters, to obtain the input guess value, i.e. $\boldsymbol{p}^0$. The quality level of a priori information depends on the standard deviation σ^p of the added $\boldsymbol{\varepsilon}^p$ error. Four different quality levels of a priori information have been considered and reported in Tab. I. The data set labeled as *guess mean* contains guesses with lowest quality and has been simulated simply by taking as a guess, the mean value of the range of each parameter. Whereas, the data sets labeled as *pb30, pb20* and *pb10* contain guesses of increasing quality. More precisely, the σ^p values have been set equal to the 30%, 20% and 10% of the whole variability range of each parameter, as summarized in Tab. II. It follows that the highest quality of a priori information is represented by the data set *pb10*, because it contains values closer to the expected parameter.

Given a (noisy) measurement and using a guess, it is expected that the optimization algorithm, after some iterations, can fit the measurement and then returns an estimated parameter closer to the true one. It follows that the error in terms of discrepancy between the estimated and expected parameters reduces from an initial value (computed on the guesses) to a final value (computed on the output parameter returned by the algorithm). The initial rms errors associated to *pb30, pb20* and *pb10* are given by the respective σ^p values shown in Tab. II. Whereas, those associated to the data set *guess mean,* are: *s*: rms=0.45 cm; *l*: rms=5.48 cm; *m*: rms=0.03; ε^r*(real)* rms=5.05; and, $m_v\%\ cm^3\ cm^{-3}$ rms=9.28.

In order to better understand the need of introducing a priori information into soil moisture retrieval algorithms, we firstly present in Fig. 1 the behavior of soil moisture estimates obtained by minimizing the cost function (6). The diagonal terms of the covariance matrix **S** have been set equal to the variance of the noise $\boldsymbol{\varepsilon}^{m\phi}$. As guess values, the data set *guess mean*, i.e. the mean values of the parameter range, has been used. This basically corresponds to the retrieval of soil moisture from backscattering values without any a priori information.

TABLE I. GUESS DATA SETS SIMULATED WITH DIFFERENT QUALITY LEVELS ADOPTED IN THE RETRIEVAL ALGORITHM AS A PRIORI INFORMATION. *EP*=EXPECTED PARAMETER VALUE, *PR*=PARAMETER RANGE, η_K=ZERO MEAN GAUSSIAN RANDOM VARIABLE, σ^r_K=STANDARD DEVIATION OF η_K.

Data set	Guess	Perturbation
guess mean	Mean value of range	-
pb30	$EP\pm\eta_3$	σ^r_3=30% *PR*
pb20	$EP\pm\eta_2$	σ^r_2=20% *PR*
pb10	$EP\pm\eta_1$	σ^r_1=10% *PR*

TABLE II. STANDARD DEVIATION σ^p USED TO SIMULATE THE GUESS FOR EACH PARAMETER OF THE IEM AND GO MODELS. THE σ^p VALUE OF THE $m_v\%$ PARAMETER IS ESTIMATED FROM ε^r BY THE EMPIRICAL RELATIONSHIP.

Data set	*s (cm)*	*l (cm)*	*m*	ε^r *(real)*	*(m_v% cm^3cm^{-3})*
pb30	σ^p=0.45	σ^p=5.40	σ^p=0.03	σ^p=5.10	(σ^p=9.60)
pb20	σ^p =0.30	σ^p=3.60	σ^p=0.02	σ^p=3.40	(σ^p=6.80)
pb10	σ^p=0.15	σ^p=1.80	σ^p=0.01	σ^p=1.70	(σ^p=3.50)

Fig. 1 shows the rms error of $m_v\%$ cm^3 cm^{-3} (i.e. $\Delta mv\%$) as a function of the standard deviation σ^ε of the noise $\boldsymbol{\varepsilon}^{m\phi}$ for different SAR data configurations (VV 23°, HH&VV 23°, HH&VV 23°&45°). As can be seen, the $\Delta mv\%$ error increases as a function of σ^ε and decreases going from the single-parameter configuration (i.e. ERS, VV 23°) to the multi-parameter configuration (i.e. HH&VV, 23°&45°).

When σ^ε is equal to zero, there is considerable difference between $\Delta mv\%$ errors of the different SAR configurations. However, as soon as σ^ε increases, the performances of the three configurations tend to get closer and closer. A strong sensitivity of $\Delta mv\%$ to σ^ε is observed. In particular, starting from $\sigma^\varepsilon \approx 1$ dB the ERS and the ENVISAT configurations are characterized by almost the same performances in terms of $\Delta mv\%$ error. Since it is unrealistic to have a total budget error significantly smaller than 1.0 dB, a priori information needs to be assimilated into the retrieval algorithm to benefit from the ASAR double-polarized configuration.

The attention is now focused on the performances of the retrieval algorithm using a priori information obtained by minimizing the functional in (5). The diagonal terms of the covariance matrices **S** and **G** and have been set equal to the variance of the noise $\boldsymbol{\varepsilon}^{m\phi}$ and $\boldsymbol{\varepsilon}^p$, respectively. The latter are the values $(\sigma^p)^2$ as reported in Tab. II.

Fig. 2 shows the $\Delta mv\%$ error of as a function of the guess quality level (related to the standard deviation σ^p) for different σ^ε values of the noise $\boldsymbol{\varepsilon}^{m\phi}$ of a typical ENVISAT configuration (i.e. HH & VV at 23° incidence angle). Moreover, the initial guess rms errors are shown by the curve plotted with the continuous line. As can be seen, the retrieval algorithm reduces significantly the initial $\Delta mv\%$ error from about 9.6 to the final error of 5.9 (at σ^ε=1.5 dB), when a priori information *pb30* is used. This rms error can be further reduced by using backscattering with a lower noise level (for example σ^ε=0.5 dB) or by using a priori information with a higher quality level (for example *pb10*). In the first case, $\Delta mv\%$ reduces up to 4.4. In the second case, it decreases to 2.8. This means that the quality level of the a priori information is the main factor that can significantly reduce the $\Delta mv\%$ error. These final $\Delta mv\%$ errors slightly increase with the noise increasing thus indicating a strong robustness of the algorithm versus σ^ε. More precisely, increasing the σ^ε up to 2.0 dB, $\Delta mv\%$ increases less than 2% for *pb30* and less than 1% for *pb10.*

In order to investigate the presence of significant bias on the retrieved soil moisture, a linear fit between expected and retrieved soil moisture values has bee carried out. Given the linear model y=A+Bx, (where x and y represent the expected and retrieved mv values, respectively) the parameter A and B have been estimated, and are: A=2.40% cm^3 cm^{-3}, B=0.92. Fig. 3 shows the scatter plot of $m_v\%$ values obtained considering the HH&VV at 23° incidence angle configuration, a noise of 1.0 dB and initial guess values correspondent to a priori information correct within 20% of the each parameter range (i.e. *pb20*). The linear fit is shown and it can be observed how the errors distribute about the fitted line. A bias is present but it is negligible.

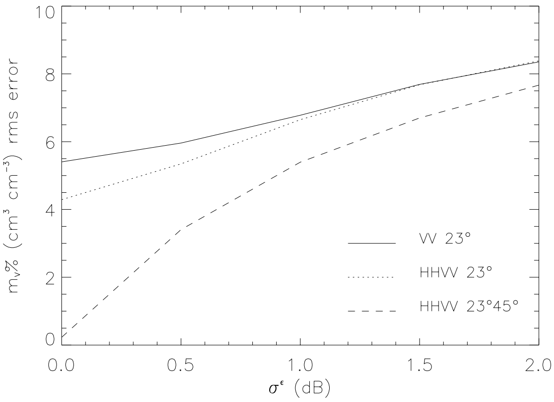

Figure 1. *m_v% rms error* versus σ^ε *noise level* at different SAR data configuration, IEM and GO simulated data, inversion without a priori information.

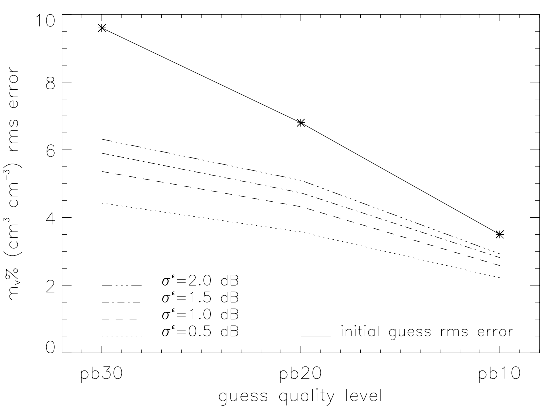

Figure 2. *m_v% rms error* versus *guess quality level* for different standard deviation σ^ε of the noise $\varepsilon^{m\phi}$, IEM and GO simulated data, HH&VV 23° SAR data configuration.

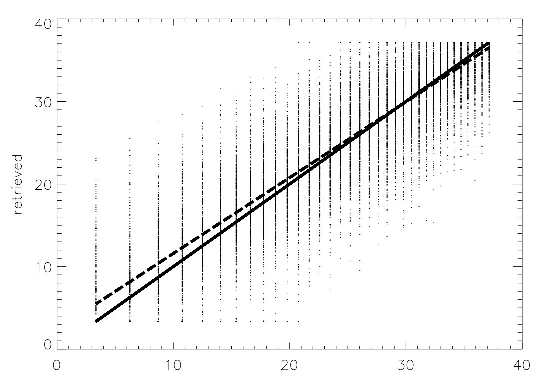

Figure 3. Scatter plot of *estimated* versus *expected m_v%.*

In summary, provided that the total error budget $\varepsilon^{m\phi}$ is less than 1.5 dB, and that the a priori information for soil moisture is correct within 20% of their whole variability range, the *soil moisture* $\Delta mv\%$ can be estimated with a rms error of about 5% $cm^3\,cm^{-3}$ by using HH&VV C-band SAR data at 23° incidence.

IV. CONCLUSIONS

In this paper, the soil moisture retrieval accuracy from ASAR measurement over soils with a large roughness variability have been investigated.

The study considered SAR data configurations corresponding to data collected by sensors on board of the ENVISAT satellite. Data were simulated by IEM and GO models, perturbed by Gaussian noise. Soil moisture content has been retrieved by using an optimization method assimilating a priori information of surface parameters.

Measurement as well as direct model errors have been considered and the retrieval performance have been estimated as a function of the a priori information on the soil surface state.

The developed algorithm demonstrates that by using a constrained optimization technique, which appropriately assimilates a priori information on soil parameters, it is feasible to retrieve soil moisture with an rms error of about 5% $cm^3\,cm^{-3}$, from HH and VV SAR backscatter at relatively small incidence angles (i.e. approximately between 20° and 35°) provided sufficiently accurate (i.e. within 20% of their whole variability range) a priori information on surface soil parameters is available.

ACKNOWLEDGMENT

This work has been partly supported by ASI contract ASI/I/R/199/02 and partly by ESA-ESTEC under contract n. 17011/03/NL/JA.

REFERENCES

[1] F. T. Ulaby, R. K. Moore, and A. K. Fung, Microwave Remote Sensing: Active and Passive. Dedham, MA: Artech House, Vol. 3, 1986.

[2] G. Satalino, F. Mattia, M. Davidson, T. Le Toan, G. Pasquariello, M. Borgeaud, "On Current Limits and Future Perspectives of Soil Moisture Retrieval from C-Band SAR data," IEEE Trans. on Geoscience and Remote Sensing, Vol. 40, No. 11, pp. 2438-2447, Nov. 2002.

[3] A.K. Fung, Microwave Scattering and Emission Models and their Applications, Artech House, Boston-London, 1994.

[4] A. Lorenc, "Analysis methods for numerical weather predictions," Quart. J. Roy. Meteor. Soc., Vol. 112, pp. 1177-1194, 1986.

[5] M. T. Hallikainen, F. T. Ulaby, M. C. Dobson, M. A. El-Rayes, and L. K. Wu, "Microwave dielectric behavior of wet soil. Part I: Empirical models and experimental observations," IEEE Trans. Geosci. Remote Sensing., Vol. GE-23, No. 1, Jan. 1985.

[6] L. S. Lasdon, A. D. Waren, A. Jain, and M. Ratner, "Design and Testing of a Generalized Reduced Gradient Code for Nonlinear Programming," ACM Transactions on Mathematical Software, Vol. 4, No. 1, pp. 34-50, March 1978.